Class 33: Time to Work on Projects

Held: Monday, 21 April 2008

Summary: We spend time exploring the kinds of analyses you should do for the project.

Notes:

- Exam 2 returned. We will discuss it on Wednesday, if you wish.
- Makeup exam Thursday a.m. One-hour block of your choice. You will get the better of (a) makeup grade, capped at 86; (b) average of makeup grade and exam grade; (c) exam 2 grade. E.g., 50 on exam 2, 82 on makeup yields 82; 80 on exam 2, 100 on makeup yields 90.
- Exam 3a given in class.
- Exam 3b is a take-home question, due Wednesday.
- Due: Mini-Project Memos.

Overview:

- Q&A.
- An example analysis.
- Time to work.

Time for Questions on the Poster

- Are there questions on the poster?

A Sample Analysis

- For many problems, you will do a chi-square style analysis.
- We’ll use as an example variables 1 (sex) and 3 (what choice was Grinnell?)

```r
> SC = fys[,c(1,3)]
> table(SC)

     CHOICE
SEX 1 2 3 4
1   4 11 34 108
2  15   7 35 127

> SCframe = data.frame(row.names=c("Below Third","Third","Second","First"),
+   Male=SCTable[1,],
+   Female=SCTable[2,])
> SCframe

          Male Female
Below Third 4     15
Third        11    7
Second       34    35
First       108   127
```
We can simply apply the `chisq.test` procedure to this table to get the important values.

```r
> chisq.test(SCframe)
   Pearson's Chi-squared test
data:  SCframe
X-squared = 6.7122, df = 3, p-value = 0.08166
```

However, we will often want to look more carefully at the differences between observed and expected values. The computation of the expected values is a strange formula that I don’t expect you to understand. (I do expect that you could do an individual computation by hand, but this does all of them at once.)

```r
> SCexpected = rowSums(SCframe) %o% colSums(SCframe)/sum(SCframe)
```

<table>
<thead>
<tr>
<th></th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Below Third</td>
<td>8.7478006</td>
<td>10.25220</td>
</tr>
<tr>
<td>Third</td>
<td>8.28739</td>
<td>9.71261</td>
</tr>
<tr>
<td>Second</td>
<td>31.76833</td>
<td>37.23167</td>
</tr>
<tr>
<td>First</td>
<td>108.19648</td>
<td>126.80352</td>
</tr>
</tbody>
</table>

We can now compare directly.

```r
> SCframe - SCexpected
```

<table>
<thead>
<tr>
<th></th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Below Third</td>
<td>-4.7478006</td>
<td>4.7478006</td>
</tr>
<tr>
<td>Third</td>
<td>2.7126100</td>
<td>-2.7126100</td>
</tr>
<tr>
<td>Second</td>
<td>2.2316716</td>
<td>-2.2316716</td>
</tr>
<tr>
<td>First</td>
<td>-0.1964809</td>
<td>0.1964809</td>
</tr>
</tbody>
</table>

As importantly, we can compute the deviations.

```r
> (SCframe-SCexpected)^2/SCexpected
```

<table>
<thead>
<tr>
<th></th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>Below Third</td>
<td>2.5768317632</td>
<td>2.1987097110</td>
</tr>
<tr>
<td>Third</td>
<td>0.8878854292</td>
<td>0.7575978934</td>
</tr>
<tr>
<td>Second</td>
<td>0.1567711671</td>
<td>0.1337667023</td>
</tr>
<tr>
<td>First</td>
<td>0.0003568024</td>
<td>0.0003044455</td>
</tr>
</tbody>
</table>

What does this tell us?

**Time To Work**

- Work on your projects.
- Cassie and I will wander around and try to answer questions.